COMPREHENSIVE BRIEF ON OPEN ACCESS TO PUBLICATIONS AND RESEARCH DATA FOR THE FEDERAL GRANTING AGENCIES

Commissioned by the Canadian Institutes of Health Research,
the Natural Sciences and Engineering Research Council of Canada and
the Social Sciences and Humanities Research Council of Canada

June 2011

Author: Kathleen Shearer, Consultant

For further information, please contact:

Canadian Institutes of Health Research
160 Elgin Street, 9th Floor
Ottawa, Ontario K1A 0W9
www.cihr-irsc.gc.ca

Natural Sciences and Engineering Research Council of Canada
350 Albert Street
Ottawa, Ontario K1A 1H5
www.nserc-crsng.gc.ca

Social Sciences and Humanities Research Council of Canada
350 Albert Street
P.O. Box 1610
Ottawa, Ontario K1P 6G4
www.sshrc-crsh.gc.ca

Available on the Web in PDF format
June 2011

The opinions expressed are those of the consultant, and do not necessarily reflect the views of CIHR, NSERC or SSHRC.

TABLE OF CONTENTS

List of Acronyms 1

Foreword 3

Executive Summary 5

I. Publications 7

1.1 Introduction 7
1.2 Policy Environment 8
1.3 Typical Policy Elements 9
1.4 Canadian Context 10
1.5 Implementation 11
1.6 Disciplinary Differences 12
1.7 International Models 12
1.8 Perspectives of Stakeholder Communities 18
1.9 Relationships with Other Policies 22
1.10 Challenges for Policy Implementation 23

II. Research Data 27

2.1 Introduction 27
2.2 Policy Environment 28
2.3 Canadian Context 29
2.4 Implementation 31
2.5 International Models 32
2.6 Perspectives of Stakeholder Communities 36
2.7 Challenges for Policy Implementation 37

III. Conclusions 40

LIST OF ACRONYMS

AAP: Association of American Publishers

ACS: American Chemical Society

ANDS: Australian National Data Service

CARL: Canadian Association of Research Libraries

CAURA: Canadian Association of University Research Administrators

CERN: European Organization for Nuclear Research

CFHSS: Canadian Federation for the Humanities and Social Sciences

CFI: Canada Foundation for Innovation

CIHR: Canadian Institutes of Health Research

CISTI: Canada Institute for Scientific and Technical Information

CRIS: Current Research Information System

DANS: Data Archiving and Networked Services (Netherlands)

DCC: Digital Curation Centre (UK)

DARIAH: Digital Research Infrastructure for the Arts and Humanities

EU: European Union

FP7: European Commission Seventh Framework Program

FRPAA: Federal Research Public Access Act (US)

IPY: International Polar Year

NARCIS: National Academic Research and Collaborations Information System

NIH: National Institutes of Health (US)

NRC: National Research Council

NSERC: Natural Sciences and Engineering Research Council of Canada

NSF: National Science Foundation (US)

OA: Open access

OCI: Office for Cyberinfrastructure (NSF)

OECD: Organization for Economic Cooperation and Development

PAGSE: Partnership Group for Science and Engineering

PIPEDA: Personal Information Protection and Electronic Documents Act

PMC Canada: PubMed Central Canada

RCUK: Research Councils UK

RDSWG: Research Data Strategy Working Group

REBs: Research Ethics Boards

SCOAP3: Sponsoring Consortium for Open Access Publishing in Particle Physics

SPARC: Scholarly Publishing and Academic Resources Coalition

SSH: Social sciences and humanities

SSHRC: Social Sciences and Humanities Research Council of Canada

STEM: Science, technology, engineering and mathematics

TCPS: Tri-Council Policy Statement on the Ethical Conduct for Research Involving Humans

FOREWORD

Canada's major public funding agencies make investments in research and research training for the benefit of all Canadians, and the world. New knowledge and insights gained through research provide solutions to many of the issues most important to Canadians: to improve the quality of our environment and health; enhance public safety and security; develop sound public policies; understand human experience and the complexity of our relations across cultures, languages, religions and histories; protect endangered species; advance economic prosperity; and so on. As such, research agencies have a fundamental interest in ensuring that the results of the research they fund are disseminated as widely as possible.

Canada's research funding agencies have been involved in open access and open data activities to varying degrees. The Canadian Institutes of Health Research (CIHR), the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Social Sciences and Humanities Research Council of Canada (SSHRC) have been monitoring open access and its implications with respect to their own research communities, and assessing the feasibility of implementing more specific policies and practices. The ultimate goal of the three agencies is to develop and implement joint policies on access to research results and data management requirements, wherever possible.

This briefing paper provides an up-to-date review of open access in order to assist the agencies in moving forward with their policy development, both individually and across the three agencies. The review outlines key recent developments in Canada and abroad with respect to sister agencies in the US, Europe, and Australia, and discusses specific challenges for Canadian granting agencies and the communities they serve. It also addresses the role of post-secondary institutions in the implementation of agency policies on open access and the potential barriers (social, cultural, structural, infrastructure, cost, etc) for the agencies in moving towards the implementation of a joint open access policy.

Although the underlying principles of open access to publications and data are similar, there are significant differences in policy elements, implementation requirements and challenges. Therefore, the paper has been divided into two sections: publications and data.

EXECUTIVE SUMMARY

Sharing and openness are the hallmarks of the scholarly tradition. Researchers publish their results, not for financial return, but to enable other researchers to build upon them and to contribute to the progress of knowledge in their fields. The Internet has fundamentally changed the practical and economic realities of distributing the results of research. Throughout the world, funding agencies, institutions, and others in the research community have turned to open access as a means to more widely disseminate the results of the research they fund and ensure that greater benefits are derived from that research.

Open access (OA) is a movement in the scholarly community to provide free and unrestricted access to the products of research. Greater access to research results is expected to accelerate the progress of research, democratize access to knowledge world-wide, and ensure that publicly funded research is available to the public.

There is a growing number of examples that illustrate how open access to research data and publications have contributed to advances in knowledge. Open access publications are used and cited more widely by other researchers. In relation to the wider community, open access contributes to the 'informed citizen' and 'informed consumer'. In addition, there is a growing body of evidence that open access would result in significant economic benefits at the national level as well.[1]

Over the last decade the momentum for open access has been steadily growing. Numerous funding agencies and institutions across the world have implemented policies requiring that the publications and data resulting from the research they support be made freely available.

Despite initial objections by publishers and some members of the research community, there is a sense that open access is here to stay and stakeholders are now looking at practical ways of adapting to this new imperative. Many Canadian organizations including the Association of Universities and Colleges of Canada (AUCC), the Canadian Association of Research Libraries (CARL), Canadian Association of University Research Administrators (CAURA), Canadian Association of Learned Journals (CALJ), Canadian Federation for Humanities and Social Sciences (CFHSS) and others are monitoring open access developments closely.

Other jurisdictions, including Europe, the U.S. and Australia, have made significant progress in terms of adopting policies and investing in infrastructure to support open access. While Canada does have a fair number of open access policies for research publications, these are primarily in the health sector and other fields have been slower to adopt such policies. With the exception of a few specific initiatives, Canadian organizations (funding agencies and institutions) have not yet implemented data sharing policies that are broadly applicable and that can be effectively monitored.

Funding agencies are faced with a number of challenges relating to open access and data sharing policies in Canada. New incentives, infrastructure, expertise, and funding models are needed, and developing these elements will require close collaboration amongst research

---

1 Knowledge Exchange comparative report on Costs and Benefits of Open Access www.knowledge-exchange.info/Default.aspx?ID=316

communities, universities, libraries, funders, and publishers. Cooperative approaches can help mitigate the risks for all stakeholders in transitioning to open access models.

In terms of publications, many of the pieces are in place in order for funding agencies to successfully implement open access policies. However, there are significant disciplinary variations in attitudes, infrastructure and support for open access to research publications, which may lead to operational differences in implementing open access policies across agencies. While there are issues in terms of publisher permissions and journal sustainability, new models are evolving rapidly and it is expected that these issues will be adequately addressed in the coming years. Potential approaches may include using disciplinary repositories (for example, PubMed Central Canada), as well as working with universities to support deposit into institutional repositories; and/or working with the Canadian publishing community to provide open access directly via open access journals.

In terms of research data, the benefits of enhancing stewardship and access are expected to be significant. However, the necessary infrastructure does not yet exist across the spectrum of disciplines and research fields to support the implementation of comprehensive data sharing policies. Most significantly, Canada does not have a comprehensive network of data repositories, nor do researchers in many fields have access to the appropriate expertise and training in data management. Nonetheless, there may be common elements for data sharing policies that can be considered. Existing models include requirements to include data management plans as part of the research proposal (as currently being implemented by the National Science Foundation in the U.S. and the International Polar Year project here in Canada), or working with universities to ensure that researchers retain their data for a given amount of time following the end of a given research project.

I. PUBLICATIONS

1.1 Introduction

Sharing and openness are the hallmarks of the scholarly tradition. Researchers publish their results, not for financial return, but to enable other researchers to build upon them and to contribute to the progress of knowledge in their fields. The Internet has fundamentally changed the practical and economic realities of distributing the results of research. Throughout the world, funding agencies, institutions, and others in the research community have turned to open access as a means to more widely disseminate the results of the research they fund and ensure that greater benefits are derived from that research.

Open access (OA) is a movement in the scholarly community to provide free and unrestricted access to the products of research.

According to the three original, formal definitions of open access (referred to as Budapest, Berlin, Bethesda; or BBB definitions), OA is "immediate, free availability on the public Internet, permitting any users to read, download, copy, distribute, print, search or link to the full text of these articles, crawl them for indexing, pass them as data to software or use them for any other lawful purpose"[2]. Open access strives not only to provide free access to the products of research, but also the ability for others to re-use and re-distribute these research results as long as there is proper attribution for the author.

The concept of open access to publications first emerged out of a meeting organized by the Open Society Institute in Budapest in 2002. The purpose of the meeting was to discuss how best to "accelerate progress in the international effort to make research articles in all academic fields freely available on the Internet."[3] This meeting was followed by a number of other public statements and declarations in support of open access by groups of scholars, funding agencies and libraries.

Since 2002, the momentum for open access has been growing, with an increasing number of research funding agencies, institutions, and research projects implementing open access policies. While they differ in their details, these policies generally require that affiliated researchers make their research articles freely available over the Internet within a given time period after publishing an article.

The momentum for open access is being driven by the conviction that greater access to research results will accelerate the progress of research, democratize access to knowledge world-wide, and ensure that publicly funded research is available to the public.

The statement published after the Budapest meeting articulates the values underlying open access:

> An old tradition and a new technology have converged to make possible an unprecedented public good. The old tradition is the willingness of scientists and scholars to publish the fruits of their research in scholarly journals without payment, for the sake of inquiry and knowledge. The new technology is the Internet. The public good they make possible is the

---

2 Budapest Open Access Initiative. www.soros.org/openaccess/read.shtml

3 Ibid

world-wide electronic distribution of the peer-reviewed journal literature and completely free and unrestricted access to it by all scientists, scholars, teachers, students, and other curious minds. Removing access barriers to this literature will accelerate research, enrich education, share the learning of the rich with the poor and the poor with the rich, make this literature as useful as it can be, and lay the foundation for uniting humanity in a common intellectual conversation and quest for knowledge.[4]

There are two principal operational means for achieving open access:

1) Open access journals are journals that do not charge readers for access. They publish their content online for free and cover their costs in other ways (such as author-pays and hybrid models). Open access journals operate like subscription-based journals in every other way, including managing the peer review process. **Open access journals are often referred to as the "gold" road to open access.**
2) Open Access repositories are databases of articles (and other materials) that are freely available to readers. An open access repository (also referred to as open access archive) may be either discipline-specific, such as PubMed Central (for the health sciences) or arXiv (which includes physics, mathematics, astronomy, computer science, quantitative biology and statistics). Alternatively, a repository may be institution-based and collect research output from all disciplines. Institutional repositories are digital collections of the outputs created within a university or research institution. **Open access repositories are often referred to as the "green" road to open access.**

There are also other variations in terms of how open access is implemented. John Willinsky from the Public Knowledge Project at UBC describes nine types of OA. "Delayed OA, for example, makes published articles freely available after a certain amount of time under subscription (typically from 3 to 24 months), and "Partial OA makes some articles from a journal issue freely available and other articles available through subscription only.[5]

In terms of open access to research publications, the main focus has been on the peer-reviewed journal literature, for which authors seek no financial compensation. However, there are a growing number of publishers offering open access monographs and this may become more wide-spread in the near future. This briefing paper will primarily concentrate on journal publications.

1.2 Policy Environment

Over the past 10 years, there have been a growing number of open access policies implemented by funding agencies around the world. SHERPA-JULIET, a service that monitors the number and type of OA policies of research funding agencies, now lists 56 agencies with open access policies, including the European Commission and the European Research Council, the US National Institutes of Health (NIH), the Norwegian Research Council, the Swiss Research Council, the Wellcome Trust, and the European Organization for Nuclear Research (CERN). According to the SHERPA-JULIET database, about 60% of existing OA mandates are based in Europe.[6]

---

[4] Budapest Open Access Initiative. www.soros.org/openaccess/read.shtml

[5] John Willinsky. "The Nine Flavours of Open Access Scholarly Publishing" www.jpgmonline.com/text.asp?2003/49/3/263/1146

[6] SHERPA-JULIET. www.sherpa.ac.uk/juliet/

In addition to funding agency policies, there are a growing number of universities and research centres that are introducing open access policies.

1.3 Typical Policy Elements

Open access policies typically require that researchers make their peer-reviewed journal articles freely available to the public via an open access repository or open access journal, often after an embargo period of from 6 to 12 months. While the specifics of policies differ according to discipline and jurisdiction, they usually include the following elements:

- **Method:** There are a variety of means by which an article is required to be made open access, such as via disciplinary and/or institutional repository, and/or open access journal.

  For example, the US National Institutes of Health and the UK Wellcome Trust both require that researchers make their articles available through PubMed Central (PMC), an open access database of biomedical literature. Other agencies, such as CIHR are less prescriptive about how articles should be made open access leaving it to the individual researcher to decide if they will deposit into an institutional or discipline based repository, or publish in an open access journal.

- **Embargo Period:** The time period following publication after which the article must be made freely available varies according to policy-usually from 6 months to 1 year, but may be longer in some fields.

  For example, some agencies require that researchers make their articles available within a given time period after publication (i.e. 6 to 12 months), while other policies simply state, "as soon as possible". The purpose of the embargo period is to protect publisher's revenue. Embargo periods differ across agencies and disciplines because it generally understood that the demand for journal articles drops off more slowly after publication in some disciplines, such as in the humanities and arts. This affects whether a journal would lose subscribers and revenue by offering open access after an embargo period of a certain length.[7]

- **Version:** Some policies require that a specific version of the article must be made available, usually either the author's manuscript or publisher's version.

- **Mandatory vs. voluntary:** Some policies are voluntary policies that request or encourage researchers to make their work Open Access, while other policies are mandatory policies that require it.

  Voluntary policies, while initially popular with funding agencies, were found to have fairly low rates of compliance. For example, the NIH policy on open access had only 4% compliance rates when it was voluntary, and has jumped to 60% when it was changed to a mandatory policy. As a result, most policies that have been recently introduced are mandatory.

- **Exceptions:** Some policies include an exception for researchers who publish in journals that do not have policies that support open access.

---

7 Promoting Open Access in the Humanities. http://www.earlham.edu/~peters/writing/apa.htm

Some open access policies, such as the CIHR Policy on Access to Research Outputs, contain an exception for researchers who are publishing in journals that do not allow open access archiving or do not offer an open access option.

1.4 Canadian Context

Canada already has a number of research agencies with open access policies. Most of these are in the health sector, and include CIHR, Canadian Breast Cancer Research Alliance, Canadian Cancer Society, Canadian Health Services Research Foundation, Fonds de la recherche en santé du Québec, Genome Canada, International Development Research Centre, and the National Research Council (NRC).

Many of the open access policies in Canada mirror the "CIHR Policy on Access to Research Outputs" which states that "Grant recipients are now required to make every effort to ensure that their peer-reviewed publications are freely accessible through the Publisher's website (Option #1) or an online repository as soon as possible and in any event within six months of publication (Option #2)."

In 2004, SSHRC adopted, in principle, a policy of open access that would guide the development of its research support programs. "Following consultations with the research community, SSHRC's governing council decided in 2006 to take an awareness-raising, educational and promotional approach to the implementation of this policy, rather than imposing mandatory requirements."[8] SSHRC has been working to promote open access through a number of projects that support the transition of Canadian humanities and social sciences journals to open access models.

In addition to these activities, CIHR, NSERC and SSHRC have developed and adopted a set of guiding principles in support of open access and have also committed to developing a shared approach for improving access to publicly funded research.

Genome Canada's Policy on Access to Research Publications states that "peer reviewed publications that have been supported, in whole or in part, by Genome Canada be made freely accessible online, in a central or institutional repository, as soon as possible, and, at the latest, six months after the publication date."[9] In support of this policy, recommendations are made encouraging researchers to publish in open access journals and journals which allow self-archiving. Genome Canada states that they will "keep this policy under review and work with other research funders to promote best practice in this area."

Canada has one university, Concordia University, which has an open access requirement for faculty. In April 2010, the Concordia University Senate passed a resolution on open access that requires its faculty members to deposit a copy of their published articles into the university's institutional repository. Athabasca University has had a voluntary Open Access Research Policy since 2006.

---

[8] SSHRC Policies: Open Access. www.sshrc-crsh.gc.ca/about-au_sujet/policies-politiques/open_access-libre_acces/index-eng.aspx

[9] Genome Canada: Policy on Access to Research Publications. http://www.genomecanada.ca/medias/PDF/EN/AccessResearchPublicationsPolicy.pdf

1.5 Implementation

Open access policies cannot be implemented without an existing infrastructure of open access repositories or journals to support them. In the current environment, it is feasible for the majority of peer-reviewed journal articles to be made open access via one of the two options for implementing open access- open access repositories or open access journals.

Open access repositories: Authors usually transfer their copyright to publishers when they sign onto publishing agreements. Authors who wish to deposit articles in an open access repository can do so legally if the publishers transfer some rights back to authors, usually through a stated publisher policy, or sometimes through authors' amendments to publishing agreements.

According to the SHERPA-RoMEO database, a service that monitors publishers open access policies, about 65% of publishers worldwide currently have policies allowing authors to deposit a copy of their article into a disciplinary or institutional repository.[10] These types of policies are also referred to as self-archiving policies. Referred to as "green" publishers, it is generally the large commercial publishing houses that tend to have these types of policies, with the result that a fairly high proportion (as many as 90%) of journals allow authors to make their articles freely available. Self-archiving policies are less frequently found in smaller publishers. (e.g., many of the small scholarly associations that publish journals do not have policies in this regard).

For journals that do not allow authors to deposit into open access repositories, authors can try attaching an author's addendum to the author agreement. These are legal instruments that modify the publisher's agreement allowing authors to keep the rights to their article, including the right to make the article open access. There are a number of these types of addenda available for authors including the SPARC Canadian Author Addendum[11].

In Canada, there are a large number of repositories available for researchers to make their articles open access. Most of the large academic libraries have implemented an institutional repository for this purpose[12], and CISTI-CIHR have partnered to create a mirror of PubMed Central, called PubMed Central Canada[13], which is available for CIHR-funded researchers to deposit their papers. There are also a number of international discipline-based open access repositories available in selected fields such as physics, mathematics, computer science, and economics.

Open access journals: According to the Directory of Open Access Journals, there are now over 6,000 Open Access journals worldwide (representing 20% of the estimated 30,000 peer reviewed journals).[14] These journals provide free access to the electronic copies of the articles they publish. In addition to these full open access journals, there are a growing number of subscription-based publishers, including the major publishing houses (Elsevier, Springer, Wiley, Sage), that offer authors the option of paying a fee to make their articles openly accessible. This

---

10 SHERPA-RoMEO. www.sherpa.ac.uk/romeo/

11 SPARC Canadian Author Addendum. www.carl-abrc.ca/projects/author/EngPubAgree.pdf

12 CARL List of Canadian Institutional Repositories http://www.carl-abrc.ca/projects/institutional_repositories/canadian_projects-e.html

13 PubMed Central Canada: http://pubmedcentralcanada.ca/index.html

14 Directory of Open Access Journals http://www.doaj.org/doaj?func=loadTempl&templ=100623&uiLanguage=en

so called 'hybrid model' enables publishers of traditional subscription-based journals to experiment with open access (see further discussion below under 'Sustainable Funding').

1.6 Disciplinary Differences

There are a number of important disciplinary differences in the uptake and implementation of open access. Currently, approximately 65% of agencies with open access policies are in the health sciences sector; the natural sciences and engineering funding agencies represent about 25% of the policies; and the social sciences and humanities agencies comprise about 10%.[15]

In some fields, such as physics and biomedicine, open access has been widely embraced and is being implemented through open access repositories. In the humanities and social sciences, the uptake for open access has been slower, and there has been a greater emphasis on implementing open access through OA journals, rather then depositing into open access repositories.

There are a number of reasons for the different disciplinary approaches and uptake of open access which are summarized below:

- Higher journal prices in science, technology, engineering and medicine (STEM), have made accessibility a bigger issue for those communities, driving open access implementation.
- Unlike STEM disciplines, much research in the SSH is produced by individual researchers without the support of a specific project grant who therefore do not have access to grant funds to publish in OA journals funded through publication fees.
- The demand for journal articles in the SSH drops off more slowly after publication than demand for articles in the STEM fields. This affects whether a journal would lose subscribers and revenue by offering open access after an embargo period of a certain length.[16] Because of this, SSH publishers have been more reluctant to adopt open access models.
- In the humanities, journals are not the only publishing vehicle, monographs are also prevalent.
- Many fields in SSH do not have an established tradition of paying for publication through page fees and there are few journals that levy author charges.
- In some fields, such as the health sciences/biomedical literature, governments have recognized that there are significant societal benefits to making this literature available to the public, practitioners, and other researchers, and have been involved in moving open access forward in those fields.
- In some fields, such as physics, there has been long tradition of sharing preprints, a way of exchanging information about research without the time lag inherent in traditional publishing.

1.7 International Models

There are several national and international examples that may act as useful models for Canada in terms of policy implementation and support for open access. The models presented here

---

15 SHERPA-JULIET www.sherpa.ac.uk/juliet/

16 Promoting Open Access in the Humanities http://www.earlham.edu/~peters/writing/apa.htm

represent a variety of approaches to implementing open access through policy adoption, legislative channels, and infrastructure development.

1.7.1 European Union

Most countries in the EU have one or more funding agencies with an open access policy, and Europe has been very active in developing national repository networks to support open access.

EU countries have benefited from two European Commission Seventh Framework Program (FP7) projects, DRIVER and DRIVER II[17], which funded the establishment and development of a European open access repository infrastructure. The projects provided funding at the national level to implement repositories, support for national help desks that provide expertise to repository developers, and also the development of a centralized search portal. The project ended in 2009, and the central portal (called DRIVER Search Portal[18]) is now being maintained collectively by national partners. It currently provides free access to over 3,000,000 research publications from 287 repositories in 38 countries.

In 2009, the European Commission (EC) began a pilot project to assess the feasibility of open access to the research funded through the FP7[19] program. The pilot aims to provide free access to peer-reviewed journal articles from FP7 funded projects. The project requires that researchers from certain fields, representing about 15% of the research funded through the FP7, are to make their articles freely available after an embargo period of 6 or 12 months, either via an open access repository or an open access journal. The pilot project will run until 2013, and the EC has said that the results will be used in their deliberations on the next steps—to improve access to research data—at both the European and national levels.[20] The EC has also funded the OpenAIRE project to build support structures for researchers in depositing FP7 research publications through the establishment of the European Help desk and the outreach to all European member states through the operation and collaboration of 27 national open access liaison offices.[21]

1.7.2 Netherlands

Dutch funding agencies have not implemented policies that require open access to their funded research; however, the country has created a robust national network of institutional repositories. The repositories were developed with funding from a national funding agency, SURF, and form part of an integrated Dutch research information system called NARCIS[22]. NARCIS offers a central access point to open access publications from the repositories of all the Dutch universities and other research institutions. Through NARCIS, open access articles are integrated with researcher bios, descriptions of research projects, and some data sets in the fields of arts and humanities, and social sciences.

NARCIS is reflective of a broader trend in Europe whereby open access repositories are being increasingly integrated with other types of research information systems, most commonly Current

---

17 DRIVER II.www.driver-repository.eu/

18 DRIVER Search. http://search.driver.research-infrastructures.eu/

19 European Commission FP7. http://cordis.europa.eu/fp7/home_en.html

20 European Commission: Open Access in FP7. http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1300

21 Open AIRE. www.openaire.eu/en/home

22 National Academic Research and Collaborations Information System. www.narcis.nl/

Research Information Systems (CRIS), systems that are aimed at gathering and disseminating data about research.

1.7.3 United Kingdom

In June 2006, the Research Councils of the UK (RCUK), an umbrella organization for the seven UK federal funding agencies, published a set of guiding principles stating that "publicly funded research must be made available to the public and remain accessible for future generations".

In recognition of different disciplinary needs, the RCUK left it to the individual funding agencies to develop policies appropriate for their given communities. So far, 6 of the 7 councils have adopted mandatory open access policies. The Engineering and Physical Sciences Research Council (EPSRC), the only council that has not yet implemented a policy, have stated that they will be developing a policy that will require open access of journal publications, "but that academics should be able to choose whether they use the so-called green option (ie, open access archiving in an on-line repository) or to use the gold option (ie, pay-to-publish in an open access journal)"[23].The Wellcome Trust, one of the worlds largest charity funders of biomedical research, has been a very strong proponent of open access and has also implemented an open access policy.

In May 2011, RCUK and the Higher Education Funding Council for England (HEFCE) announced a joint commitment to open access. Their public statement sets out the principles of how they will work together:

> "Research Councils UK and HEFCE have a shared commitment to maintaining and improving the capacity of the UK research base to undertake research activity of world leading quality, and to ensuring that significant outputs from this activity are made available as widely as possible both within and beyond the research community. Open access to published research supports this commitment and, if widely implemented, can benefit the research base, higher education, and the UK economy and society more broadly. To achieve this, open access needs to be implemented with clear licensing agreements, sustainable business models, and working with the grain of established research cultures and practices.
>
> HEFCE and the Research Councils will work together and with other interested bodies to support a managed transition to open access over the medium term, and welcome the work of the UK Open Access Implementation Group in support of this aim."[24]

UK open access policies are supported by a comprehensive network of institutional repositories managed by the universities, that were developed with some funding from a central funding agency, the Joint Information Systems Committee. Because not all universities currently have an institutional repository, the UK has also set up central repository, called The Depot, in which authors from all disciplines can deposit their papers in order to comply with open access policies. They have also developed a UK PubMed Central, in which biomedical researchers, funded through Wellcome Trust and the UK Medical Research Council, must make their articles available.

---

[23] Engineering and Physical Sciences Research Council: Policy on access to research outputs. www.epsrc.ac.uk/about/infoaccess/Pages/roaccess.aspx

[24] Research Councils UK and HEFCE joint commitment on open access. http://www.rcuk.ac.uk/research/Pages/outputs.aspx

1.7.4 Australia

The two national funding agencies in Australia (Australian Research Council and the National Health and Medical Research Council) both have policies that encourage recipients of grants to make their published work openly accessible, and asks that grantees produce reasons in their final grant report if this did not happen. About 50% of Australian Universities have institutional repositories to collect research papers and make them freely available.

While open access to publications has not been a priority for Australian funding agencies, access to data has been a key strategic aim. In 2008, the Federal Department of Industry, Innovation, Science & Research provided initial funding of 48 million over two years to establish the Australian National Data Service. This initiative is described in detail in the data section of this report.

1.7.5 United States

In the US, open access is being pursued through legislative channels. In 2007, a mandatory provision for NIH-funded research papers was included in a bill passed by Congress and approved by the President. This bill required that the NIH implement an open access policy. The policy was put in place in 2008 and states that "all investigators funded by the NIH submit or have submitted for them to the National Library of Medicine's PubMed Central an electronic version of their final, peer-reviewed manuscripts upon acceptance for publication, to be made publicly available no later than 12 months after the official date of publication"[25].

To expand on the NIH legislation, the Federal Research Public Access Act (FRPAA) was introduced in 2006 and again in 2010 with bi-partisan support in Congress. The proposed bill would require federal agencies with annual extramural research budgets of $100 million or more to provide the public with online access to research manuscripts stemming from funded research no later than six months after publication in a peer-reviewed journal. The bill would require agencies, such as the National Science Foundation (NSF) (which currently has no policy on access to publications) to implement open access policies similar to the one enacted by the NIH. The bill has yet to be made into law, and must now be reintroduced for a third time. The bill's prospects for reintroduction may be limited because it might not be among their top priorities in the new Congress. If passed, the bill would have a significant impact on the Canadian landscape given the large number of researchers in Canada that receive funding from US funding agencies.

Another possible option for OA policy implementation in the US lies in a potential executive policy implementation. In 2010, the Obama administration collected comments on the potential implementation of a broadened NIH-style mandate that would cover any federal agency with an extramural budget in excess of $100 million. However, after having collected the responses the administration has yet to act.[26]

There are varying levels of infrastructure support for open access across the US. There are several large disciplinary repositories housed in the United States, including PubMed Central and arXiv. Many US universities have an institutional repository, but there has not been any

---

[25] Revised Policy on Enhancing Public Access to Archived Publications Resulting from NIH-Funded Research. http://grants.nih.gov/grants/guide/notice-files/NOT-OD-08-033.html

[26] Hadro, Josh. As COMPETES Act Is Signed into Law, 'Wait-and-See' Is the Attitude on Further OA Legislation. Library Journal. Jan 20, 2011. www.libraryjournal.com/lj/home/888910-264/as_competes_act_is_signed.html.csp

centralized funding to support the development and maintenance of repositories and repository implementation varies significantly across the country.

1.7.8 Summary Table of International Open Access Models

Jurisdiction	European Commission	Australia	Canada	Netherlands	United Kingdom	United States
Legislation	No	No	No	No	No	Law requiring NIH funded researchers to make articles open access. Draft FRPAA law to be introduced in Congress
Policies	Yes (pilot project involving 15% of FP7 research)	No	Yes, mainly health and SSH funding agencies	No	Yes all national councils with the exception of EPSRC	National Institutes of Health
Infrastructure Support	Institutional repositories developed through national program and FP7 DRIVER	Universities developing OA repositories	PubMed Central Canada (for health research). All major Canadian research universities have repositories, but no funded national program. Some support for OA journals through Erudit and Synergies projects	Institutional repositories Content integrated into a national CRIS	UK PMC, The Depot (a national central repository), and institutional repositories	PMC and institutional repositories (for health research)
Other support programs	OpenAire project to coordinate deposit of articles across participating nations	No	No	National funding program to develop and populate repositories; Funding from EC DRIVER project	National funding program to develop and populate repositories; Funding from EC DRIVER project	No

1.8 Perspectives of Stakeholder Communities

Numerous consultations and surveys over the last ten years demonstrate that most stakeholder communities involved in scholarly research are supportive of open access in principle, but may have operational concerns related to their specific perspective.

1.8.1 Researchers and Students

Open access is not a concept widely implemented in the research community; however, there has been a growing awareness among researchers over the last several years. Studies have found that researchers' attitudes towards open access are generally positive, but they have concerns about how it will impact them in terms of their funding and their freedom to publish where they choose.[27,28]

In Canada, both CIHR and SSHRC have undertaken consultations with their research communities about open access. In 2005, SSHRC staff conducted a survey across a significant range of actors, including researchers and scholars, scholarly associations, publishers, editors and librarians to elicit comments and views on the subject of open access. Of the 130 respondents, most expressed their support for open access, although many had concerns with the financial issues for small journal publishers in the transition to open access.[29]

In 2006, CIHR posted a draft "Policy on Access to Research Outputs" and launched a consultation process to gather feedback. They received about 150 submissions. They found that most researchers favoured open access publishing over open access repositories as the means for providing free access to peer-reviewed publications. In addition, many researchers and publishers requested that CIHR provide additional grant funds to cover the costs of article-processing fees charged by publishers for making articles open access.

More recently, an international survey conducted in 2010 of over 40,000 researchers from across disciplines found that there was "overwhelming support for the idea of open access, while highlighting funding and (perceived) quality as the main barriers to publishing in open access journals"[30]. This study looked at researchers' attitudes towards open access publishing (not open access repositories) and found that 89% of respondents considered open access publishing to be beneficial for their research field. The survey also found that funding was the major barrier to publishing in open access journals, followed by the presence of journals of a (perceived) suitable quality.

Major concerns

In summary, the major concerns about open access expressed by the research community are described below:

---

27 Swan, Alma and Shridan Brown. Open Access Self Archiving: An Author Study. May 2005. www.jisc.ac.uk/uploaded_documents/Open%20Access%20Self%20Archiving-an%20author%20study.pdf

28 Creaser, Claire , Fry, Jenny , Greenwood, Helen , Oppenheim, Charles , Probets, Steve , Spezi, Valérie and White, Sonya 'Authors' Awareness and Attitudes Toward Open Access Repositories', New Review of Academic Librarianship, 16:1, 145 - 161.

29 Chan, Leslie, Frances Groen, and Jean-Claude Guédon. Feasibility of Open Access Publishing for Journals Funded by the Social Science and Humanities Research Council of Canada, pg. 2.

30 Dallmeier-Tiessen, Highlights from the SOAP project survey. What Scientists Think about Open Access Publishing. Jan 28, 2011. http://arxiv.org/abs/1101.5260

- Researchers often feel that open access material is of lower quality than subscription-based journal content.
- Copyright issues and the lack of clarity of publishers' open access policies are perceived as a major barrier to open access archiving. Generally, authors do not know whether they can upload a copy of their article onto a repository or website, nor do they know which version (pre-print, author's final copy or publisher's copy) they can mount.
- The deposit process itself is also often cited as a barrier. The time it takes and technical know-how to deposit articles into repositories contribute to authors' concerns about depositing material in an open access repository.
- Researchers are concerned about the financial sustainability of scholarly publishing in an open access environment. In particular, many of the scholarly societies (of which researchers are members) rely on the revenues generated through journal subscriptions to fund parts of their operations.
- Researchers feel that open access requirements will infringe on their freedom to publish where they feel is most appropriate.
- Researchers have no means to pay open access publication fees.

1.8.2 Research Institutions

A growing number of universities world-wide have adopted open access policies, including MIT, Harvard, and Stanford in the US, as well as institutions in Australia, Belgium, Finland, Germany, India, Italy, Norway, Portugal, Spain, Switzerland, Turkey, and the UK. These university-wide or departmental policies are typically implemented by faculty through a resolution or vote, and call for faculty to deposit their peer-reviewed articles into the university repository.

In Canada, several organizations have already expressed interest and/or support for open access. Open access has been the topic of discussion at meetings and conferences for many organizations over the last several years. For example, the Canadian Association of University Research Administrators will have Open Access as a program item at their annual general meeting in mid-May 2011.

In 2000, the Association of University and Colleges of Canada (AUCC) expressed support for the Scholarly Publishing and Academic Resources Coalition (SPARC), a US-based organization advocating for open access on behalf of the US and Canadian library communities. Since then, however, the organization has made no public comments about the issue.

In terms of individual universities, Concordia University is the first Canadian University to adopt a requirement for faculty to make their articles available via the open access repository at the university. In April 2010, the Concordia University senate passed a Resolution on Open Access, which states,

"Any scholarly article accepted for publication in a peer-reviewed journal, from now on requires all faculty members to deposit an electronic copy in *Spectrum* along with non-exclusive permission to preserve and freely disseminate it. This requirement is not binding in cases where publishers, co-authors or other rights holders disallow such a deposit. Faculty members may also opt out of the requirement by notifying the University Librarian in writing that their work has

appeared, or will appear in another Open Access format; or by citing other factors that currently discourage them from depositing their work in an Open Access repository."[31]

Athabasca University has had an Open Access Research Policy since 2006. The policy is voluntary and asks "faculty, academic and professional staff deposit an electronic copy of any published research articles (as elsewhere accepted for publication) in an AU repository. The contract with the publisher determines whether the article is restricted (lives in the repository as a record of the AU's research but is not accessible online by searchers) or open access (accessible online by searchers)."[32]

1.8.3 International Publishing Community

Some resistance to open access has come from commercial and scholarly society publishers who are concerned about the viability of economic models in an open access environment. The most prominent issues put forward by the publishing community are the issues of financial sustainability of open access models. In particular, publishers argue that:

- Open access archiving will reduce subscription revenues to the point where scholarly societies cannot continue to exist. Memberships will be cancelled without a subscription-based journal, as a tangible benefit to membership in the society.
- Open access publishing offers no means for publishers to recoup their costs (or make profits).

While most publishers publicly support open access in principle, they are often opposed to obligatory requirements by funding agencies and institutions, and in the case where there are requirements, they call for funding to be provided for authors to pay for publication fees that may be charged by publishers.

Two of the most vocal opponents to open access have been the American Chemical Society (ACS), the world's largest scientific society, and the Association of American Publishers (AAP). The ACS position is stated on their website: "The American Chemical Society (ACS) supports universal access to the results of scientific research via publishing models that are sustainable and that ensure the integrity and permanence of the scholarly record upon which scientific progress is based. The ACS does not support unfunded mandates that place constraints on authors or that interfere with our ability to fulfill the Society's mission as a provider of indispensable information to the world's community of chemistry professionals."[33]

The Association of American Publishers argues that, "Policies that mandate open access publishing unilaterally force scientists to limit themselves to open-access journals or hybrid journals or risk violating the agreements they have with their publishers. Scientists should not be limited to publishing in a few compliant journals. Doing so limits intellectual freedom and scholarly independence and is, quite simply, against the public interest. Scientists and their publishers understand and support the government's goal to broaden the accessibility of research, and they have incentives and are committed to making research widely available.

---

[31] Concordia University Senate Resolution on Open Access: Approved April 16, 2010. www.library.concordia.ca/research/openaccess/SenateResolutiononOpenAccess.pdf

[32] Athabasca University: Open access research policy. www.athabascau.ca/policy/research/openaccess.htm

[33] American Chemical Society: Ensuring Access to High Quality Science. http://portal.acs.org/portal/PublicWebSite/policy/publicpolicies/balance/highqualityscience/WPCP_011535

However, forcing publishers to adopt a singular business model that might not be appropriate is not supported by sound economic policy."[34]

1.8.4 Canadian Publishers

Canada has a relatively small academic publishing community and funding is a major challenge for Canadian journal publishers. Most Canadian-based journals are run on a shoestring budget, and rely heavily on volunteer contributions, graduate student work, and technical and hosting support from academic libraries. A number of Canadian journals are subsidized through government programs such as the SSHRC Aid to Scholarly Journals Program; however, the subsidies do not cover the full costs of running the journals.

Similar to the international publishing community, the major concern for Canadian publishers in adopting open access is finding a sustainable business model. In 2010, the Association of Canadian University Presses produced a White Paper on Open Access. The paper states "the sustainability of an open access based business model is a key concern for scholarly publishers who are considering offering OA products"[35]. There are however, a number of Canadian Presses that have begun to adopt some open access publishing options such as Athabasca University Press and the University of Calgary Press.[36] The Canadian Association of Learned Journals has been monitoring developments with open access, and has been actively looking for sustainable models for implementing open access in member journals.

Several existing Canadian publishing initiatives in Canada provide support for open access. The NRC Research Press, which publishes 17 journals in the sciences, allows authors to deposit articles into open access repositories 6-months after they are published. In September 2010, the NRC Research Press transitioned from the federal government into an independent not-for-profit organization, and is introducing a pay for open access option. Érudit, a multi-institutional publishing consortium (l'Université de Montréal, de l'Université Laval et de l'Université du Québec à Montréal), which hosts over 150 journals and also publishes the journals supported by the Fonds québécois de recherche sur la société et la culture, requires that all journals it hosts provide open access to their publications within two-years of publication. And, Synergies Canada, a collaborative initiative of twenty-one Canadian universities funded through the Canada Foundation for Innovation (CFI) to transition Canadian SSH journals from print to electronic format, is working with 170 Canadian journals, many of which are open access. In addition, there are a number of independently run scholarly journals that are open access, such as *Analyses : Revue de Critique et de Théorie Littéraire*, *Open Medicine* and *Advances in Science*.

1.8.5 Canadian Academic Associations

A number of Canadian associations representing different stakeholder communities have made public statements about open access. In March 2006, the Canadian Federation for Humanities and Social Sciences (CFHSS) published a position statement on open access. The statement expressed support for open access in principle, but recommended an incremental transition to OA, without mandates imposed by funding agencies or universities. Their key concern with open access was the financial viability of scholarly societies. However, despite these concerns, open access was one of the themes of the Congress 2010 which was held at Concordia University in Montreal.

---

34 What is "open access". www.publishers.org/issues/5/8/

35 Kwan, Andrea. Open Access and Canadian University Presses. Associations of Canadian University Presses. 2010. pg. 3

36 Shearer, Kathleen. A Review of Emerging Models in Canadian Academic Publishing. University of British Columbia. 2010. https://circle.ubc.ca/handle/2429/24008

The Canadian Association of University Teachers has stated its support for the concept of open access. Their focus has been on the intellectual property aspect of open access and they strongly advocate for authors to retain the copyright for their work, rather than signing it over to publishers as is the standard requirement by subscription-based publishers.

The National Graduate Caucus of the Canadian Federation of Students, which represents over 70,000 graduate students from across the country, has officially endorsed open access. Other student groups from campuses across Canada have also come out in favour of open access.

The Canadian Association of Research Libraries (CARL) has been a strong advocate for open access in Canada. The association has actively lobbied governments to require funding agencies and universities to implement open access policies. Through a number of projects, CARL has also been providing support for members to set-up institutional repositories, promote open access on campus, and host open access journals.

1.9 Relationships with Other Policies

As a growing number of funding agencies and universities adopt open access policies, there may be potential issues for authors who are funded by more than one agency or are affiliated with an institution that also has an open access policy. Possible areas of conflict may include prescribed method of deposit, length of embargo period, and version of paper to be made open access. For example, a funding agency policy may require that an article be made open access in a disciplinary repository (e.g. NIH requires deposit into PubMed Central), while a university requires that an article be made available via the university's institutional repository. The potential problems arising from these conflicts are not insurmountable, and are being addressed through the development of technologies (e.g. that facilitate dual deposit) or through harmonization of policies.

There are a number of potential conflicts between publishers' policies and funding agencies open access policies. While typical funding agency policies require that the authors' final manuscript be made available within 6 to 12 months of publication, there are still a number of publishers (approximately 37%[37]) that do not allow their articles to be made publicly available; and still others that only allow open access to the "pre-print" copy of the article, rather than the authors' final manuscript. In addition, some publishers may have embargo periods of up to two years, which are often longer than those imposed by funding agency policies.

CIHR and some other institutions have addressed the issue of conflict with publisher policies by including an "opt out" option for authors who are publishing in journals that have conflicting policies. Others, such as the NIH, simply require grantee compliance regardless of publisher policy. Authors must publish elsewhere if the publisher refuses to accommodate the NIH policy.

Promotion and tenure processes/policies of universities, while not in direct conflict with open access policies, can work against them. Promotion and tenure criteria intrinsically deny recognition to new journal publications, many of which may be open access, and often act to deter submissions to them. In terms of open access repositories, many prestigious journals allow authors to archive their articles; however, institutions rarely take into account authors' efforts to deposit into open access repositories as a criterion in their promotion and tenure processes. In

[37] SHERPA-RoMEO. www.sherpa.ac.uk/romeo/statistics.php

some cases, there is the perception that some journal articles in repositories are not peer reviewed.

1.10 Challenges for Policy Implementation

There are a number of important issues that Canada's funding agencies may want to consider before implementing open access policies.

1.10.1 Operational Feasibility

The vast majority of researchers in Canada currently have access to one or more of the options available to make their articles open access. However, since not all researchers in Canada have an open access repository at their institution, or will want to publish in an open access journal, there will be some who will not be in compliance with an open access policy.

This problem exists in all jurisdictions and has been addressed by others in different ways. The NIH, Wellcome Trust, and several of the UK funding councils, simply require compliance with the policy and insist that authors publish only in those journals compliant with their policies. Other agencies, such as CIHR, allow authors to opt out of the open access requirement if publishers do not allow open access archiving. The CIHR policy states, "Publications must be freely accessible within six months of publication, where allowable and in accordance with publisher policies."[38] Such opt out clauses do result in lower rates of open access to articles and therefore agencies need to assess how they will impact the availability of the articles resulting from the research they fund.

Other organizations are working to provide support for open access infrastructure that will help as many researchers as possible to comply with policies. In Europe, several governments, including the Netherlands, Germany, UK, and others have invested heavily in strengthening repository networks to ensure OA repositories are available to all researchers. In addition, individual organizations or funding agencies, such as the European Commission through the EC Research Framework, the Wellcome Trust as well as a number of universities have set up funds for authors to pay publishing fees for making articles open access.

1.10.2 Sustainable Funding

Open access policies can only be effective in an environment where there is a sustainable repository infrastructure and/or open access journals. Both these options present some inherent challenges in terms of funding.

Open access journals

Open access publishing essentially requires a re-distribution of funds from a subscription based model, whereby users or their libraries pay for access to articles, to other models whereby publishers recoup their publishing costs and profits in other ways (e.g., authors' fees). Many publishers are adopting new business models to support open access.

According to the Directory of Open Access Journals, about 20% of the peer-reviewed journals world-wide are already open access.[39] These journals employ a number of different business models which include subsidies, advertising, charges for hard copy versions, charges for other

---

38 CIHR Policy on Access to Research Outputs. http://www.cihr-irsc.gc.ca/e/34846.html

39 Directory of Open Access Journals. http://www.doaj.org

publication services, membership fees,[40] or some combination of these. There is, however, a clear trend towards a publication fee model that requires authors to pay to publish their articles in an open access journal. This is also the model most often used by hybrid publishers that offer a paid open access option. The fees vary widely and range from $75 to $3500 per article[41], depending on the journal.

Current arrangements for paying open access fees in Canada and internationally have grown up haphazardly and are not standardized. Generally speaking, across the world, most OA publishing fees are currently being paid by funding agencies. A few academic libraries are now providing access to funds set aside specifically for open access fees. In some cases, however, the costs are met from unallocated funds from research grants or other sources.

This pay to publish model will have implications for funding agencies that wish to support their researchers in publishing in open access journals. Some funding agencies have set up special funds to pay for authors who wish to publish in journals that charge open access fees. Many agencies in the sciences already have dedicated funding for page charges and have adapted that to include open access journal fees. The Wellcome Trust, for example, provides grant holders with additional funding, through their institutions, to cover open access charges in order to meet the Trust's open access requirements. Other organizations, such as the Max Planck Society in Germany and the European Commission provide authors with full reimbursements for the cost of publishing in an open access journal.

In Canada, open access publication fees are listed as an eligible expense for the dissemination of research results in the Tri-Agency Financial Administration Guide.[42] Canada also has some libraries with special funds to pay for authors at their institution who wish to publish in journals that charge open access fees (Simon Fraser University, the University of Calgary, and the University of Ottawa). The libraries then request that publishers decrease the subscription price slightly for every open access article for which they pay."[43]

Some journals, such as BioMed Central, charge open access fees and also offer institutional memberships, which then reduce or eliminate the per article fees they charge. In these scenarios, fee structures for institutional memberships vary from publisher to publisher, but fees are usually tiered based on the size of the institution.

Publishers are also experimenting with other business models. In 2009/2010, several universities in US and Europe have entered into agreements with Springer whereby articles written by affiliated authors will be made fully and immediately open access for a flat fee paid by the institution. In this case, there are no separate per-article charges, since costs have been factored into the overall fee. However, anecdotally it is said that Springer will not be entering into any

---

[40] Some journals that charge article-processing fees offer institutional memberships which then reduce or eliminate the per article fees they charge. The fee structures for institutional memberships vary from publisher to publisher, but fees are usually tiered based on the size of the institution; with smaller institutions paying less than large ones.

[41] Open Access Directory: Publication Fees http://oad.simmons.edu/oadwiki/OA_journal_business_models#Publication_fees

[42] 2010 Tri-Agency Financial Administration Guide http://www.nserc-crsng.gc.ca/professors-professeurs/financialadminguide-guideadminfinancier/index_eng.asp

[43] University of Calgary Open Access Fund: Frequently Asked Questions. http://library.ucalgary.ca/services/for-faculty/open-access-authors-fund/open-access-authors-fund-frequently-asked-questions-faq

more of these type agreements in the future, likely because they do not consider it to be a sustainable model for the future if broadened to other institutions.

Collaborative agreements between publishers, libraries and funding agencies are thought to be one way of mitigating the risks of transitioning to open access. There are already a number of collaborative projects attempting to transition journals to open access models:

- The SCOAP3 (Sponsoring Consortium for Open Access Publishing in Particle Physics) project is a group of High Energy Physics (HEP) funding agencies and research libraries that are coordinating to cover journal subscription prices so that publishers can make the electronic versions of their journals freely available over the Internet. There are no article processing fees, and authors are not charged directly to publish their articles. The project is spearheaded by CERN, with partners in Canada, France, Germany, Italy, Sweden and the US. Each SCOAP3 partner will finance its contribution through the cancellation of journal subscriptions for the 7 core peer reviewed journals in HEP. Project partners have estimated that the total amount of money currently spent by the library community on these 7 titles worldwide is about $15M US.[44] The project has now garnered enough support from countries world-wide in order to begin moving ahead.
- In Canada, the Synergies Project is a CFI funded collaborative initiative involving publishers and libraries that is assisting Canadian SSH journals move from print to electronic formats, and many of these journals have adopted open access models.

Open access repositories

Sustainable funding is also an issue for repositories and other associated open access infrastructures (such as harvesters or portals that provide access to the content of a group of repositories). For now, most of the costs of running institutional repositories are being covered by the universities, usually from the library operational budgets. However, since repositories are not yet considered to be central to the operations of most libraries, institutional repositories can be vulnerable to funding cuts.

There are also few mature and sustainable funding models for disciplinary repositories. Traditionally, disciplinary archives have been maintained by government agencies or universities, but as the content in repositories grows in volume, they are becoming more expensive to maintain. It is unlikely that many funding agencies will want to adopt the centralized model, like NIH and CIHR, in which the agency provides funding for the repository (see section 4 above).

New models for funding repositories are also being sought. The physics preprint ArXiv, for example, which was co-funded by Cornell University and the NSF until 2010, is now seeking to broaden its sources of funding. ArXiv has been asking individual universities and research organizations to make annual voluntary contributions based on the amount of downloading utilization by each institution.

Currently in Canada, CIHR and CISTI are providing ongoing funding for PMC Canada and the universities are funding institutional repositories. However, current funding levels for institutional repositories are relatively small and there are few funding programs to support the development of aggregation services, which could be an important tool for monitoring and analyzing researcher compliance.

---

[44] SCOAP3 http://scoap3.org/index.html

1.10.3 Researcher Awareness

While awareness of open access is growing, its implementation is still not widespread in the research community. In addition there are a number of common misperceptions held by researchers about open access.

Numerous surveys over the last decade have found that many researchers have a confused understanding of the 'open access' concept, its purpose and the means by which to achieve it.[45,46,47,48] For example, many researchers think that open access can only be achieved by publishing in an open access journal and are unaware of the "green" road to open access through repositories. In addition, many researchers are not aware that their institution has an open access repository, or that most publishers allow them to deposit their articles into the repository.

Some agencies that have open access policies have addressed this issue by providing detailed information to researchers discussing the benefits of open access, and addressing some of the common misperceptions that might act to hinder the effective implementation of an open access policy. Agencies could work with other organizations in Canada that already support open access, such as CARL, the Canadian Library Association, the Canadian Medical Libraries Association, and CFHSS to raise awareness of the issue across Canada.

1.10.4 Compliance and Enforcement

Few OA policies currently include strong sanctions for researchers who are non-compliant and it is unclear how many agencies are monitoring compliance. Some agencies have stated that non-compliance may impact future funding decisions. Both the NIH and the Wellcome Trust have tried to address non-compliance of their open access policies by sending letters to grantees, reminding them of their obligation under the funding agreement. These letters were also sent to grantees' institutions.

For agencies using a central repository as the locus of OA material, such as NIH and CIHR, monitoring compliance is relatively easy. The NIH policy, for example, requires grantees to use the "Manuscript Submission reference number" in future progress reports or funding applications. Grantees obtain the reference numbers when they deposit their work in PMC, meaning agencies can monitor compliance fairly easily. For policies that do not mandate a central repository as the mode for open access, tracking open access articles in disparate repositories and publishers' websites has the potential to become an onerous and time consuming process.

Ideally, as with some other funding agency policies, the universities could play a role in monitoring compliance with policies, but in practice, it is not clear how this could be implemented.

---

[45] Swan, Alma and Shridan Brown. Open Access Self Archiving: An Author Study. May 2005. www.jisc.ac.uk/uploaded_documents/Open%20Access%20Self%20Archiving-an%20author%20study.pdf

[46] Creaser, Claire , Fry, Jenny , Greenwood, Helen , Oppenheim, Charles , Probets, Steve , Spezi, Valérie and White, Sonya 'Authors' Awareness and Attitudes Toward Open Access Repositories', New Review of Academic Librarianship, 16:1, 145 - 161.

[47] Chan, Leslie, Frances Groen, and Jean-Claude Guédon. 2006. Feasibility of Open Access Publishing for Journals Funded by the Social Science and Humanities Research Council of Canada.

[48] Dallmeier-Tiessen, Highlights from the SOAP project survey. What Scientists Think about Open Access Publishing. Jan 28, 2011. http://arxiv.org/abs/1101.5260

II. RESEARCH DATA

2.1 Introduction

For the purpose of this briefing paper, research data are defined as the factual records used as primary sources for research, and that are commonly accepted in the research community as necessary to validate research findings.

As with open access to publications, there is a broad international trend towards open access to data. This trend, often referred to as 'data sharing' is reflected in numerous international reports published over the last decade that have called for a greater sharing of research data within and across disciplines. These reports assert that improving access to research data would have significant benefits for research and society such as: accelerating scientific progress, avoiding the duplication of research, enabling replication and verification of research results, and increasing the visibility and impact of research.

Most recently, in October 2010, a High Level Expert Group submitted a report to the European Commission that described a vision for a pan-European data infrastructure, with links to the international community. The report "identifies the benefits and costs of accelerating the development of a fully functional e-infrastructure for scientific data - a system already emerging piecemeal and spontaneously across the globe, but now in need of a far-seeing, global framework. The outcome will be a vital scientific asset: flexible, reliable, efficient, cross-disciplinary and cross-border."[49]

Several Canadian consultations over the past decade have discussed the potential benefits of data sharing in Canada:

- In October 2002, SSHRC and the National Archivist of Canada established a Working Group that recommended the creation of a new national research data archival service.
- In November 2004, NRC, in partnership with CFI, CIHR and NSERC, undertook a National Consultation on Access to Scientific Research Data (NCASRD) in the natural and medical sciences community. The final report provides a "road map" for the implementation of a national plan for open access to publicly funded scientific research data.
- In March 2010, a report entitled *Canadian Digital Information Strategy: Final Report of Consultations with Stakeholder Communities 2005-2008* published by Library and Archives Canada after extensive consultations with organizations across Canada called for greater sharing and preservation of research data within governments and the research community.
- In 2010, the Canadian government published a Digital Economy Consultation Paper. The paper says, "Governments can help by making publicly-funded research data more readily available to Canadian researchers and businesses. Open access is consistent with many national strategies and holds great economic potential for Canadians to add value to machine-readable data, while ensuring that privacy rights are protected. In many cases, data are already available but are difficult to locate. Consistent methods of access will be reinforced."[50]

---

[49] Riding the Wave. http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf

[50] Digital Economy Consultation: Consultation Paper. http://de-en.gc.ca/consultation-paper/consultation-paper-6/

Canada has not yet put into practice the recommendations from these various reports.

In 2008, a multi-stakeholder group, called the Research Data Strategy Working Group (RDSWG), was formed to addresses the challenges and issues surrounding the access and preservation of data arising from Canadian research. The Working Group includes representatives from universities, data centers, research libraries and CIOs, the granting agencies, government science departments and agencies, Compute Canada, and the research community. Their activities focus on the actions and leadership roles that organizations can take to ensure Canada's research data is accessible and usable for current and future generations of researchers. The Working Group is currently planning a Research Data Summit for senior policy makers and university administrators in order to develop a roadmap for more comprehensively managing research data in Canada.

The trend towards greater data sharing in the research environment is developing in parallel with a trend towards "open data". Open data initiatives are aimed at expanding access and creative use of government-generated data into the non-governmental sphere by encouraging innovative ideas, tools and web applications. In March 2011, the Government of Canada launched the Open Data pilot project, an online data portal that provides access to a large number of government datasets through a single window. The data can be reused by application developers for commercial or research purposes[51].

2.2 Policy Environment

Data sharing policies are most often developed and implemented by research funding agencies and in some cases research projects, and less commonly adopted by universities.[52] While data sharing policies differ across organizations, typical policy elements go beyond asking researchers to retain research data for a given period of time and include more comprehensive requirements to ensure that data are both retained and available to others.

Requirements for data sharing can range from full public open access, to sharing with specific researchers upon request, to access governed through restrictive licenses, depending on the sensitivity of the data, the size and complexity of the data set, their perceived reuse value, and the availability of a repository.

Typical policy elements of data sharing policies are described below:

- **Data management plans:** Investigators are required to submit a data management plan with their funding proposals. These plans ensure that researchers consider ahead of time how they will manage and share their data.
- **Data quality and standards:** Investigators are required to adhere to international standards that will ensure the data is accessible by others.
- **Data documentation:** Data documentation and metadata must accompany data so that the data is understandable by others.
- **Method of data sharing:** Investigators are required to either (1) deposit data in relevant subject or institutional repositories; or, (2) where there are no repositories hold the data locally, and make it available through a web-based presence; or (3) retain data so that upon request, other researchers can have access to data.
- **Timing of data sharing:** Investigators must make data accessible within a given period of time after publication of research results.

---

51 Open Data Pilot Project. http://www.data.gc.ca/default.asp?lang=En

52 SHERPA-JULIET. www.sherpa.ac.uk/juliet/

- **Data retention:** Data should be retained for a minimum number of years (on average 5 years)
- **Data preservation:** Investigators must deposit their data in a long-term repository, where available, to ensure the preservation of their data.

There are also a number of common exceptions that are often included in data sharing policies:

- **Privacy and confidentiality of data:** The privacy of individuals who participate in research and the confidentiality of the data must be protected at all times. Data intended for broader use must be free of identifiers that would permit linkages to individual research participants and variables that could lead to deductive disclosure of the identity of individual participants. In some cases where data cannot be stripped of identifiers, for example longitudinal studies that collect data over a period of time and must compare data points, data may be exempted from the data sharing requirements or data sharing may be qualified.
- **Intellectual property:** Policies may permit delays in sharing research data for a period of time, in cases whereby institutions or researchers are applying for patents or developing new applications based on that data.
- **Traditional knowledge:** Where local and traditional knowledge is concerned, rights of the knowledge holders shall not be compromised.
- **Sensitive data:** Where data release may cause harm, specific aspects of the data may need to be kept protected (for example, locations of nests of endangered birds or locations of sacred sites, or data related to national security)

2.3 Canadian Context

In Canada, as elsewhere, data sharing practices are very discipline specific. In some fields- such as genomics, proteomics, high-energy physics, and astronomy- data archiving and sharing is the norm. In other fields no such traditions exist. There is, however, a growing awareness across the scholarly community that there are significant potential benefits in making research data available for re-use.

In 2004, Canada along with 33 other countries (including the US, China, Japan and many European countries) adopted the OECD Declaration on Access to Research Data From Public Funding[53]. The underlying principles of this declaration are that publicly-funded research data are a public good, produced in the public interest, and that they should be openly available to the maximum extent possible.

These same principles have been reflected in other discipline-based initiatives. CIHR, for example, has recently signed a joint statement with the Wellcome Trust, the National Institutes of Health and other health funding agencies. It is a statement of intent to improve data sharing and reads, "we, as funders of health research, intend to work together to increase the availability to the scientific community of the research data we fund that is collected from

[53] OECD Declaration on Access to Research Data from Public Funding. www.oecd.org/document/0,2340,en_2649_34487_25998799_1_1_1_1,00.html

populations for the purpose of health research, and to promote the efficient use of those data to accelerate improvements in public health."[54]

Both SSHRC and CIHR have policies related to research data. SSHRC's Research Data Archiving Policy has been in place since 1990. The policy states that "All research data collected with the use of SSHRC funds must be preserved and made available for use by others within a reasonable period of time. SSHRC considers "a reasonable period" to be within two years of the completion of the research project for which the data was collected." There are few mechanisms in place, such as data repositories and data management expertise, to support researchers in preserving their data and there is no oversight regarding the implementation of this policy.

As part of its broader policy on access to research outputs, CIHR requires grant recipients to deposit certain data types—bioinformatics, atomic, and molecular coordinate data—into the appropriate public database immediately upon publication of research results. CIHR also requires researchers to retain original data sets arising from CIHR-funded research for a minimum of five years after the end of the grant. CIHR has indicated that they will review and update this policy on an annual basis or as needed.[55]

Projects funded by Genome Canada must comply with its Policy on Data Release and Resource Sharing, with expectations to share data and resources as rapidly as possible. At a minimum, data is expected to be released and shared "no later than the original publication date of the main findings from any datasets generated by that project."[56] At the completion of a project, all data is to be shared without restriction. In addition, applicants must submit a Data and Resource Sharing Plan with each funding application. Genome Canada has an additional policy on Intellectual Property[57] to ensure the proper management of acquired data and resources.

There are also an increasing number of data sharing policies at the level of the research project. The NEPTUNE Project, an underwater ocean observatory at the University of Victoria, which makes huge volumes of data openly available to the public, has a Data Access Policy. The International Polar Year (IPY), a large scientific program focused on the Arctic and the Antarctic from March 2007 to March 2009, had a comprehensive data policy which "requires that IPY data, including operational data delivered in real time, are made available fully, freely, openly, and on the shortest feasible timescale."[58] Dozens of Canadian research projects were selected for IPY funding from a variety of sources including the federal government, territorial governments, granting agencies and foundations.

There are also several other policies in Canada governing the management of research data:

- The 2nd edition of the Tri-Council Policy Statement on the Ethical Conduct for Research Involving Humans (TCPS) sets out privacy and confidentiality requirements for researchers working with human participants, including for secondary use of research data. The policy

---

54 Sharing research data to improve public health: full joint statement by funders of health research. www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/Public-health-and-epidemiology/WTDV030690.htm

55 CIHR Policy on Access to Research Outputs. www.cihr-irsc.gc.ca/e/34846.html

56 Genome Canada: Policy on Data Release and Resource Sharing. http://www.genomecanada.ca/medias/PDF/EN/DataReleaseandResourceSharingPolicy.pdf

57 Genome Canada: Policy on Intellectual Property. http://www.genomecanada.ca/medias/PDF/EN/IntellectualProperty.pdf

58 Canadian IPY Data Policy 2007-2008. www.ipy-api.gc.ca/pg_IPYAPI_055-eng.html

emphasizes that respect for privacy in research is an internationally recognized norm and ethical standard.

- All data in Canada collected, used or disclosed during the course of commercial activities are also subject to the Personal Information Protection and Electronic Documents Act.
- Researchers that use federal government data may also be governed by data policies and licence agreements in terms of the reuse and accessibility of their data.

2.4 Implementation

Policies that require data sharing cannot be implemented without corresponding infrastructures and other support mechanisms. Data cannot remain on the hard drives of researchers, but must be transferred to an environment where they are managed appropriately.

Ensuring the long term accessibility of research data is a complex and resource intensive process. Data must be created and maintained in a manner consistent with the goal of long-term preservation and involves active data management throughout the life-cycle of the data, beginning at the time they are first envisioned. The data must also be integrated into an enduring institutional environment supported by a stable digital repository.

There are some large scale international data repositories in certain fields, such as PubChem, GenBank, Protein Data Bank, Digital Sky Survey, World Data Centers, Global Biodiversity Information Facility, International Virtual Observatory Alliance, the Inter-university Consortium for Political and Social Research (ICPSR), and so on. These repositories collect data from around the world and provide broad access to the data in order to further research and knowledge creation. The vast majority of these archives are funded through government departments and/ or funding agencies in the country in which they are housed.

In addition, governments around the world maintain repositories that house data in many areas deemed of national importance, including climate data, population statistics and health data. The data housed in these government repositories are typically generated by governments, but are often made accessible to academic researchers for their research (though often through a pay per use option).

Similarly, Canada has a number of large government repositories and discipline-based repositories managed by universities and research centres. However, according to a Gap Analysis conducted by the Research Data Strategy Working Group in 2008, there are large gaps in both coverage and capacity of data repositories in Canada. Repositories do not exist for all subject areas, and the vast majority of research data still rests on researchers' hard drives or locked in cabinets. Only a few active data repositories in Canada allow researchers to deposit their data.[59]

Institutional repositories, based at universities, have until recently put emphasis on the deposit of textual research output (e.g., journal articles and theses). The scope of these repositories is gradually being extended to cover research data as well, but the overall number of stored datasets is still very low. While institutional data repositories hold promise for the future with the advantage of being close to researchers, they are short of expert know-how and resources. As well, the business case for supporting a data repository is not yet clear for many research institutions.

---

[59] Stewardship of Research Data in Canada: A Gap Analysis, 2008. http://dsp-psd.tpsgc.gc.ca/collection_2009/cnrc-nrc/NR16-123-2008E.pdf

To address the current lack of infrastructure in Canada, the Canadian Association of Research Libraries (CARL) is proposing to develop a national network of repositories for collecting research data, in collaboration with other partners. The vision for the project is to develop data repositories at Canada's universities in which researchers could deposit their data and link them with discipline-based repositories so that data can be integrated and reused in new ways. The project is currently in its initial stages. However, once the conceptual model has been developed, CARL will be seeking CFI funding that would enable them to lay the foundations for this project.

2.5 International Models

Few, if any, countries, currently have the infrastructure required to support widespread data sharing policies. However, several other jurisdictions are moving towards implementing the support mechanisms required to facilitate the widespread sharing and re-use of research data.

2.5.1 European Commission

In terms of data sharing policies, the EC, through the Seventh Framework Programme (FP7), requires that all research projects develop a preliminary data management plan as part of their proposals describing how data derived from the project will be managed.

The EC does not maintain data repositories, but through FP7, they have been funding e-infrastructure projects at EU member-states, including the development of discipline-based data repositories. One example of these projects is DARIAH (Digital Research Infrastructure for the Arts and Humanities)[60], which aims to enhance and support digitally-enabled research across the humanities and arts. DARIAH is developing repository infrastructure that will support of ICT-based research practices. Researchers will be able to go to DARIAH to find data and tools, archive their data, exchange information and advice in the field of metadata and digitalizing. Construction of DARIAH will begin sometime in 2011.

2.5.2 Netherlands

In the Netherlands, the Research Data Forum[61] has recently been launched in order to improve how research data is managed and to enable better access to such data for scientists/scholars and the public. The forum brings together initiatives developed by a number of different organizations and focuses on the technical, infrastructural, legal, and organizational aspects of storing research data and making it accessible.

The Royal Netherlands Academy of Arts and Sciences and the Netherlands Organisation for Scientific Research maintain the Data Archiving and Networked Services (DANS). Since its establishment in 2005, DANS has been storing and making research data in the arts, humanities and social sciences permanently accessible. DANS maintains a permanent archiving service, stimulates others to follow suit, and works closely with data managers to ensure as much data as possible is made freely available for use in scientific research. DANS is open to all researchers in the arts, humanities and social sciences in the Netherlands, and enables them to both store their data and to search for data themselves.

---

60 DARIAH. www.dariah.eu/

61 Surf Foundation. www.surffoundation.nl/en/actueel/Pages/Collaboratingonimprovedaccesstoresearchdata.aspx

2.5.3 United Kingdom

The UK has some of the most comprehensive data sharing policies of any government. Four of the seven Research Councils within RCUK have data policies in place that require their researchers to make their research data available "with as few restrictions as possible in a timely and responsible manner to the scientific community for subsequent research."[62] The policies vary, but generally researchers are also expected to make use of existing standards for data collection and management and make data available through existing community resources or databases where possible.

The UK also has a very robust infrastructure of discipline-based data repositories for collecting research data, managed by several of the RCUK funding agencies, and they have been providing centralized funding to develop university based repositories that are capable of collecting research data. The UK also has set up the Digital Curation Centre (DCC)[63], a centre of expertise for curating digital research data. In addition to providing expert advice and training to researchers in the area of data management, they are a gateway to the technical solutions, curation tools and learning resources that can help data custodians build capacity for digital curation.

2.5.4 Australia

In Australia, the funding agencies have not implemented data sharing policies but are investing heavily in infrastructure. In 2008, Australia launched a comprehensive national program for data sharing called the Australian National Data Service (ANDS)[64] as part of its National Collaborative Research Infrastructure Strategy.

The aim of ANDS is to create the infrastructure to enable Australian researchers to easily publish, discover, access and re-use research data. Their approach has been to engage in partnerships with the research institutions to support better local data management that enables structured collections to be created and published. ANDS then connects those institutional collections so that they can be found and used through the Australian Research Data Commons. The Australian Research Data Commons represents a significant change in their perspective towards research data, considering data as a strategic national resource.

Through ANDS, the Australian government is investing over 10 million dollars per year to support the development of data repositories, metadata and support services, and centralised access services through the Data Commons.

2.5.5 United States

In the US, both the National Institutes of Health (NIH) and the National Science Foundation (NSF) have policies in regards to data sharing. NIH has had a data sharing policy since 2003. The policy applies only to projects submitting a research application requesting $500,000 or more of direct costs in any single year. The policy states that "Data should be made as widely and freely available as possible while safeguarding the privacy of participants, and protecting confidential

---

62 Biotechnology and Biological Sciences Research Council: Data sharing policy. BBRSRC Data Sharing Policy. www.bbsrc.ac.uk/publications/policy/data_sharing_policy.pdf

63 Digital Curation Centre. http://www.dcc.ac.uk/

64 Australian National Data Service. http://ands.org.au/

and proprietary data. NIH investigators are also expected to include a plan for sharing final research data for research purposes, or state why data sharing is not possible."[65]

NSF has a policy on dissemination and sharing of research results that reads, "investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants."[66] They have also recently announced new requirements that all NSF proposals include a data management plan in the form of a two-page supplementary document describing how researchers will conform to the policy. According to the NSF, "This is the first step in what will be a more comprehensive approach to data policy,"[67]

In terms of infrastructure, the US is home to several large-scale discipline-based data repositories supported by the NIH, NSF and other government agencies. In 2005, the NSF instituted an Office for Cyberinfrastructure (OCI). The OCI's "Cyberinfrastructure Vision for 21st Century Discovery"[68] sets out the vision the NSF is to pursue in making research data accessible. The NSF's goals for the period of 2006-2010 are to catalyze the development of a system of science and engineering data collections that is open, extensible, and evolvable; and to support development of a new generation of tools and services for data discovery, integration, visualization, analysis and preservation. To realize this vision, NSF has provided $100 million in funding over five years for a program called "Sustainable Digital Data Preservation and Access Network Partners (DataNet)". The program is working with some of the large scale data repositories to develop "new methods, management structures and technologies to manage the diversity, size, and complexity of current and future data sets and data streams by creating a set of exemplar national and global data research infrastructure organizations".

---

65 NIH Data Sharing Policy and Implementation Guidance. http://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm

66 NSF Data Sharing Policy. www.nsf.gov/bfa/dias/policy/dmp.jsp

67 Press Release 10-077. Scientists Seeking NSF Funding Will Soon Be Required to Submit Data Management Plans. www.nsf.gov/news/news_summ.jsp?cntn_id=116928&org=NSF

68 NSF 07-28, Cyberinfrastructure Vision for 21st Century Discovery. www.nsf.gov/pubs/2007/nsf0728/index.jsp

2.5.6 Summary Table of National Data Sharing Activities

Jurisdiction	European Commission	Australia	Canada	Netherlands	United Kingdom	United States
Legislation	No	No	No	No	No	No
Policies	Data management FP8 will likely introduce stricter data sharing requirements	No	CIHR requires data deposit with certain types of data. SSHRC has a policy, but is not mandatory	No	4 of the 7 research councils have data sharing policies	Both NIH and NSF have data sharing policies
National Infrastructure Support	Discipline and national data repositories; varying levels of implementa-tion depending on country; FP7 funding development of data repositories and interopera-bility across national repositories	Large scale project to build discipline and institutional data repositories	Selective disciplinary data repositories, but not widespread; IPY Data Assembly Network; CFI support for the development of databases during the life of the research project	Large national disciplinary repositories in Humanities and Arts, Social Sciences and	Large national disciplinary data repositories attached to funding agencies; a few institutional data repositories	Selective disciplinary data repositories NSF DataNet Program to further develop disciplinary projects
Other support programs	Some projects provide discipline based support services. Varying levels of support at the national level	ANDS provides expertise; support for metadata and standards	No	DANS repository provides central expertise; support for metadata and standards in Humanities	Digital Curation Centre provides expertise; support for metadata and standards	No

2.6 Perspectives of Stakeholder Communities

There is growing recognition of the merits of data sharing in principle and in practice in the broader research community.

2.6.1 Researchers

Researchers' perspectives towards data sharing are very discipline specific. Surveys and interviews undertaken over the last decade have articulated a wide range of opinions on the topic which cannot be easily generalized into a single statement about researchers' attitudes.

Some fields have a tradition of data sharing and researchers have become comfortable with the concept. In other fields, researchers are still very opposed to making their data available for a number of reasons. Typical objections to data sharing fall in the areas of data ownership, time and skills involved with managing data, and privacy issues involving data about human participants.

A UK study of 16 different disciplines describes a number of factors that explain the differences in disciplinary attitudes[69]:

- the heritage and practices of niche research communities;
- the type and quantity of data they produce;
- the uniqueness of those data and their potential value in terms of reuse;
- the propensity of each community to create, adapt or adopt common data formats, metadata schema and other relevant standards;
- their willingness to share data in a world where competition for funding looms large;
- the policies of funding bodies in relation to data management, sharing and preservation;
- the provision of storage infrastructure including national data centres and effective discovery systems;
- the size of research teams (larger teams can benefit from keeping its own data private).

A recent review of the literature across 15 international jurisdictions undertaken in the Netherlands found that "although there are major differences in the way disciplines conduct their research, they also have a number of factors in common when it comes to data storage and access. They all encounter both technical barriers, for example the use of obsolete software, and non-technical ones, such as fear of competition, lack of trust, lack of incentives, and lack of control."[70]

One particularly important issue expressed by researchers is that they remain in control of what happens to their data. Researchers wish to control who has access to the data and under which conditions.

There is a growing awareness in the research community of the value of data sharing. This was reflected in a number of submissions to Canada's Digital Economy Consultation that called for greater government support to assist researchers in making their data available. In addition, the Partnership Group for Science and Engineering (PAGSE), for example, has called on the

---

69 Digital Curation Centre: SCARP. http://www.dcc.ac.uk/projects/scarp

70 SURF Foundation: What Researchers Want. www.surffoundation.nl/en/publicaties/Pages/Whatresearcherswant.aspx

government to "make data generated from federally funded research freely available online and provide the capacity to ensure data stewardship and preservation in the long term."[71]

2.6.2 Institutions/Universities

To date, Canadian universities have not been actively engaged in supporting researcher data sharing practices. In terms of data policies, universities enforce data privacy policies via their Research Ethics Boards (REBs) in conformity with the TCPS. They have not developed policies on data sharing and have not been enforcing compliance with the data sharing policies of funding agencies.

Regarding infrastructure requirements, some universities host and provide financial support for discipline-based databases and repositories, but this support extends to a small minority of research projects. University libraries, which currently have services that provide access to data housed elsewhere (e.g., Statistics Canada data) through research data centres, are becoming interested in collecting the research data created at their institution. However, for the most part, data management support through the university libraries is still in its infancy.

One project that may act as an important demonstrator for institutional support for data sharing policies is the IPY Data Assembly Centre Network. The network is being developed to archive and provide access to all observed data and information generated from IPY projects funded by the Government of Canada Program for IPY. The Network in its current form consists of partners from the research library community (University of Alberta, Ontario Council of University Libraries' Scholars Portal) and several government agencies. The startup funding for this project is being provided by the department of Aboriginal Affairs and Northern Development, but the ongoing expenses of managing and preserving the data into the future will be eventually taken on by individual institutions.

2.7 Challenges for Policy Implementation

There are a number of issues Canadian funding agencies may want to consider when implementing data sharing policies.

2.7.1 Skills, Training and Qualified Personnel

An important requirement for data accessibility is that data are organized and described using standards and best practices. This requires a significant amount of skill in terms of data management. A Gap Analysis[72] published by the Canadian Research Data Strategy Working Group in 2008, concluded that researchers rarely have the skills required to appropriately manage their data. The situation is similar in other countries.

As noted above, both the UK and Australia have created national centres of expertise to provide support for the research community and to data repository managers. In the United States, some university libraries have been working with researchers to assist them with managing their data appropriately. Regardless of the model, researchers in many disciplines will need access to support services in order to comply with any data sharing policy.

---

[71] PAGSE Submission to House of Commons Standing Committee on Finance 2010 Pre-budget Consultation. http://www.pagse.ca/en/briefs/sub2010e.htm

[72] Stewardship of Research Data in Canada: A Gap Analysis. http://dsp-psd.tpsgc.gc.ca/collection_2009/cnrc-nrc/NR16-123-2008E.pdf

2.7.2 Complex Policy Environment

The wide range of data policies that govern different jurisdictions and types of research data make it very challenging for researchers to understand and adhere to data sharing policies. This is particularly so for researchers who are working with data related to human participants. The Tri-Council Policy Statement on the Ethical Conduct for Research Involving Humans (TCPS) requires that data be completely anonymized or de-identified before they are shared, unless the researcher can justify to the Research Ethics Board (REB) otherwise. The 2nd edition of the TCPS provides guidance on the collection, use, dissemination, retention, and disposal of data. A narrow interpretation of TCPS by REBs or researchers can result in the unnecessary destruction of data related to human subjects in contravention with data sharing policies.

There are ways of ensuring that data sharing policies don't conflict with or compromise privacy and confidentiality requirements. The National Institutes of Health (NIH) policy on research data sharing, for example, states, "Prior to sharing, data should be redacted to strip all identifiers, and effective strategies should be adopted to minimize risks of unauthorized disclosure of personal identifiers."[73] Similarly, the Wellcome Trust policy requires the anonymization of data to protect confidentiality and insists that data confidentiality should not "unduly inhibit responsible data sharing for legitimate research uses."[74]

Other jurisdictions are developing clear instructions for researchers and REBs as to how to comply with funding agency data sharing policies in this complex environment. These could be provided in the form of "best practice" documents which offer clear guidance on how to comply with data sharing policies.

2.7.3 Infrastructure Support

For research data to be available after the lifespan of a specific research project, they must be integrated into an enduring institutional environment supported by a digital repository. The preservation of research data requires the active management of data over its entire lifecycle and involves activities such as "appraising, selecting, depositing or ingesting data into a repository, ensuring authenticity, managing the collection of data and metadata, refreshing digital media, and migrating data to new digital media."[75]

Currently in Canada, most of the data collected through research are not deposited into data repositories and few if any repositories have full preservation capacity. Although data in certain disciplines are being collected by national agencies, this represents only a small minority of data sets created through research activities in Canada.

The lack of infrastructure is most acute when looking at the hundreds of smaller datasets produced by individual researchers and research groups. It is often suggested that institutional repositories are the natural locus for such datasets. However, existing institutional repository platforms do not yet have the functionality required for data to be tagged at the element level, something that is needed for interoperability and re-use of data. In addition, because research

---

[73] NIH Data Sharing Policy and Implementation Guidance. http://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm

[74] Wellcome Trust: Guidance for researchers: Developing a data management and sharing plan. www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/Guidance-for-researchers/index.htm

[75] Stewardship of Research Data in Canada: A Gap Analysis. http://dsp-psd.tpsgc.gc.ca/collection_2009/cnrc-nrc/NR16-123-2008E.pdf

data are highly heterogeneous it is unlikely that any single repository could collect the range of data types created at any given single higher education institution.

2.7.4 Clarifying Roles and Responsibilities

There are currently large gaps in the roles and responsibilities for managing research data across its lifecycle. Researchers are responsible for managing their data during the lifespan of the project, but lack the means to maintain data once the project is over, and often lack the skills to prepare it for dissemination.

Again, institutions are an obvious candidate for taking on responsibility for curating the data produced by their own research community where those data have no natural home. This, however, would require that institutions become aware of their potential role in the management of research data. In addition, there are significant costs associated with collecting and preserving research data, and there are not yet sustainable funding models in Canada that support these activities.

III. CONCLUSIONS

Governments have a strong interest in developing efficient scientific information systems that maximize the impact of public investments in research. Improving the linkages between research and society is a key strategic aim. The Internet has made these linkages much stronger by offering the opportunity to share information and data in an unprecedented way.

There is a growing number of examples that illustrate how open sharing of data and research publications have contributed to advances in knowledge. In relation to the wider community, open access contributes to the 'informed citizen' and 'informed consumer'. A study by Australian economic researchers suggests there are substantial societal benefits that are financially measurable[76,77].

Other jurisdictions, including Europe, the U.S. and Australia, have made significant progress in terms of adopting policies and investing in infrastructure to support open access. While Canada does have a fair number of open access policies for research publications, these are primarily in the health sector and other fields have been slower to adopt such policies. With the exception of a few specific initiatives, Canadian organizations (funding agencies and institutions) have not yet implemented data sharing policies that are broadly applicable and that can be effectively monitored.

Funding agencies are faced with a number of challenges relating to open access and data sharing policies in Canada. New incentives, infrastructure, expertise, and funding models are needed, and developing these elements will require close collaboration amongst research communities, universities, libraries, funders, and publishers. Cooperative approaches can help mitigate the risks for all stakeholders in transitioning to open access models.

Currently, there are significant disciplinary variations in attitudes, infrastructure and support for open access to research publications, which may lead to operational differences in implementing open access policies across agencies. Potential approaches may include using disciplinary repositories (for example, PubMed Central Canada), as well as working with universities to support deposit into institutional repositories; and/or working with the Canadian publishing community to provide open access directly via open access journals.

The benefits of enhancing stewardship and access to research data are expected to be significant. However, the necessary infrastructure does not yet exist across the spectrum of disciplines and research fields to support the implementation of comprehensive data sharing policies. Most significantly, Canada does not have a comprehensive network of data repositories, nor do researchers in many fields have access to the appropriate expertise and training in data management. Nonetheless, there may be common elements for data sharing policies that can be considered. Existing models include requirements to include data management plans as part of the research proposal (as currently being implemented by the National Science Foundation in the

---

76 Houghton, J., Steele, C. and Sheehan, P. 2006, Research Communication Costs in Australia: Emerging Opportunities and Benefits, Report to the Department of Education, Science and Training, Canberra, CSES, Victoria University, Melbourne.

77 Houghton, J., Rasmussen, B. and Sheehan, S.; with Oppenheim, C., Morris, A., Creaser, C., Greenwood, H., Summers, M. and Gourlay, A. 2009, Economic implications of Alternative Scholarly Publishing Models: Exploring the Costs and Benefits, JISC EI-ASPM Project, Report to the Joint Information Systems Committee (JISC) (UK), CSES and Loughborough University, January.

U.S. and the International Polar Year project here in Canada), or working with universities to ensure that researchers retain their data for a given amount of time following the end of a given research project.